

## Refine Search

### Search Results -

Terms	Documents
L17 same packet	18

**Database:**

- US Pre-Grant Publication Full-Text Database
- US Patents Full-Text Database
- US OCR Full-Text Database
- EPO Abstracts Database
- JPO Abstracts Database
- Derwent World Patents Index
- IBM Technical Disclosure Bulletins

**Search:**

L18

Refine Search

Recall Test

Clear

Interrupt

### Search History

**DATE: Wednesday, November 09, 2005** [Printable Copy](#) [Create Case](#)

<u>Set Name</u>	<u>Query</u>	<u>Hit Count</u>	<u>Set Name</u>
			result set
<i>DB=PGPB,USPT,USOC,EPAB,JPAB,DWPI,TDBD; PLUR=YES; OP=OR</i>			
<u>L18</u>	L17 same packet	18	<u>L18</u>
<u>L17</u>	L4 same cluster same link	161	<u>L17</u>
<u>L16</u>	L13 same cluster	7	<u>L16</u>
<u>L15</u>	L14 same cluster	5	<u>L15</u>
<u>L14</u>	L13 same packet	49	<u>L14</u>
<u>L13</u>	L12 same (cyclic or crc)	119	<u>L13</u>
<u>L12</u>	L4 same link	8765	<u>L12</u>
<u>L11</u>	L10 same (cyclic or crc)	3	<u>L11</u>
<u>L10</u>	L8 same link	562	<u>L10</u>
<u>L9</u>	L8 same 15	0	<u>L9</u>
<u>L8</u>	multi-point	10219	<u>L8</u>
<u>L7</u>	L6 same link	13	<u>L7</u>
<u>L6</u>	L5 same 14	16	<u>L6</u>
<u>L5</u>	intra-cluster	352	<u>L5</u>

<u>L4</u>	point-to-point	30284	<u>L4</u>
<u>L3</u>	L2 same 11	8	<u>L3</u>
<u>L2</u>	inter-cluster adj1 packet	10	<u>L2</u>
<u>L1</u>	intra-cluster adj1 packet	12	<u>L1</u>

END OF SEARCH HISTORY

[First Hit](#)[Previous Doc](#)[Next Doc](#)[Go to Doc#](#)[Generate Collection](#)[Print](#)

L3: Entry 3 of 8

File: PGPB

Feb 10, 2005

DOCUMENT-IDENTIFIER: US 20050034033 A1

TITLE: Communication between and within multi-processor clusters of multi-cluster computer systems

**Abstract Paragraph:**

Improved techniques are provided for detecting and correcting errors and skew in inter-cluster communications within computer systems having a plurality of multi-processor clusters. The local nodes of each cluster include a plurality of processors and an interconnection controller. Intra-cluster links are formed between the local nodes, including the interconnection controller, within a cluster. Inter-cluster links are formed between interconnection controllers of different clusters. Intra-cluster packets may be serialized and encapsulated as inter-cluster packets for transmission on inter-cluster links, preferably with link-layer encapsulation. Each inter-cluster packet may include a sequence identifier and error information computed for that packet. Clock data may be embedded in symbols sent on each bit lane of the inter-cluster links. Copies of transmitted inter-cluster packets may be stored until an acknowledgement is received. The use of inter-cluster packets on an inter-cluster link is preferably transparent to other links and to the protocol layer.

**Summary of Invention Paragraph:**

[0010] Intra-cluster packets may be encapsulated as inter-cluster packets and serialized for transmission on inter-cluster links, preferably with link-layer encapsulation. Each inter-cluster packet may include a sequence identifier and error information computed for that packet. Clock data may be embedded in symbols sent on each bit lane of the inter-cluster links. Copies of transmitted inter-cluster packets may be stored until an acknowledgement is received. The use of inter-cluster packets on an inter-cluster link is preferably transparent to other links and to the protocol layer.

**Summary of Invention Paragraph:**

[0011] Some embodiments of the invention provide a computer system including a plurality of processor clusters interconnected by a plurality of point-to-point inter-cluster links, each processor cluster comprising nodes including a plurality of local processors and an interconnection controller interconnected by a plurality of point-to-point intra-cluster links, communications within a cluster being made via an intra-cluster protocol that uses intra-cluster packets, wherein the interconnection controller in each cluster is operable to map locally-generated communications directed to others of the clusters to the point-to-point inter-cluster links and to map remotely-generated communications directed to the local nodes to the point-to-point intra-cluster links, communications between clusters being made via an inter-cluster protocol that uses inter-cluster packets, an inter-cluster packet encapsulating at least one intra-cluster packet, each interconnection controller configured to compute a first cyclic redundancy code check for, and to encode first cyclic redundancy code check data in, each inter-cluster packet transmitted on the point-to-point inter-cluster links.

**Summary of Invention Paragraph:**

[0013] The nodes may be configured to compute a cyclic redundancy code check on a transmission window that includes transmissions of multiple intra-cluster packets

on the point-to-point intra-cluster links. Each interconnection controller may be further configured to compute a cyclic redundancy code check on a transmission window that includes transmissions of multiple intra-cluster packets on the point-to-point intra-cluster links. Each interconnection controller may encode the cyclic redundancy code check data in a field reserved for a link layer of each inter-cluster packet transmitted on the point-to-point inter-cluster links.

Summary of Invention Paragraph:

[0023] Some aspects of the invention provide a computer-implemented method for detecting errors in a computer system comprising a plurality of clusters, each cluster including a plurality of local nodes and an interconnection controller interconnected by point-to-point intra-cluster links, communications between the local nodes and the interconnection controller made via an intra-cluster protocol using intra-cluster packets, the interconnection controller of each cluster interconnected by point-to-point inter-cluster links with the interconnection controller of other clusters. The computer-implemented method includes the following steps: forming an inter-cluster packet by encapsulating an intra-cluster packet; encoding a sequence identifier in the inter-cluster packet; calculating first cyclic redundancy code check data based only upon the inter-cluster packet; encoding the first cyclic redundancy code check data in the inter-cluster packet; and transmitting the inter-cluster packet from a first interconnection controller to a second interconnection controller on a point-to-point inter-cluster link. The encoding steps may include encoding in an area of the inter-cluster packet reserved for link layer information.

Summary of Invention Paragraph:

[0025] Some embodiments of the invention provide an apparatus for detecting errors in a computer system comprising a plurality of clusters, each cluster including a plurality of local nodes and an interconnection controller interconnected by point-to-point intra-cluster links, communications between the local nodes and the interconnection controller made via an intra-cluster protocol using intra-cluster packets, the interconnection controller of each cluster interconnected by point-to-point inter-cluster links with the interconnection controller of other clusters. The apparatus includes: a device for forming an inter-cluster packet by encapsulating an intra-cluster packet; a device for encoding a sequence identifier in the inter-cluster packet; a device for calculating first cyclic redundancy code check data based only upon the inter-cluster packet; a device for encoding the first cyclic redundancy code check data in the inter-cluster packet; and a device for transmitting the inter-cluster packet from a first interconnection controller to a second interconnection controller on a point-to-point inter-cluster link.

CLAIMS:

18. A computer system comprising a plurality of processor clusters interconnected by a plurality of point-to-point inter-cluster links, each processor cluster comprising nodes including a plurality of local processors and an interconnection controller interconnected by a plurality of point-to-point intra-cluster links, communications within a cluster being made via an intra-cluster protocol that uses intra-cluster packets, wherein the interconnection controller in each cluster is operable to map locally-generated communications directed to others of the clusters to the point-to-point inter-cluster links and to map remotely-generated communications directed to the local nodes to the point-to-point intra-cluster links, communications between clusters being made via an inter-cluster protocol that uses inter-cluster packets, an inter-cluster packet encapsulating at least one intra-cluster packet, each interconnection controller configured to: receive an intra-cluster packet from the local processors via the intra-cluster links of the coherent interface; encapsulate the intra-cluster packet as a high-speed link packet by adding a sequence identifier and a packet type identifier to the intra-cluster packet; serialize the high-speed link packet; and forward the serialized high-speed link packet to the inter-cluster interface for transmission to another

processor cluster via an inter-cluster link.

[Previous Doc](#)

[Next Doc](#)

[Go to Doc#](#)

[First Hit](#)[Previous Doc](#)[Next Doc](#)[Go to Doc#](#)[Generate Collection](#)[Print](#)

L3: Entry 1 of 8

File: PGPB

Feb 10, 2005

DOCUMENT-IDENTIFIER: US 20050034049 A1

TITLE: Communication between multi-processor clusters of multi-cluster computer systems

**Abstract Paragraph:**

Improved techniques are provided for detecting and correcting errors and skew in inter-cluster communications within computer systems having a plurality of multi-processor clusters. The local nodes of each cluster include a plurality of processors and an interconnection controller. Intra-cluster links are formed between the local nodes, including the interconnection controller, within a cluster. Inter-cluster links are formed between interconnection controllers of different clusters. Intra-cluster packets may be serialized and encapsulated as inter-cluster packets for transmission on inter-cluster links, preferably with link-layer encapsulation. Each inter-cluster packet may include a sequence identifier and error information computed for that packet. Clock data may be embedded in symbols sent on each bit lane of the inter-cluster links. Copies of transmitted inter-cluster packets may be stored until an acknowledgement is received. The use of inter-cluster packets on an inter-cluster link is preferably transparent to other links and to the protocol layer.

**Summary of Invention Paragraph:**

[0010] Intra-cluster packets may be encapsulated as inter-cluster packets and serialized for transmission on inter-cluster links, preferably with link-layer encapsulation. Each inter-cluster packet may include a sequence identifier and error information computed for that packet. Clock data may be embedded in symbols sent on each bit lane of the inter-cluster links. Copies of transmitted inter-cluster packets may be stored until an acknowledgement is received. The use of inter-cluster packets on an inter-cluster link is preferably transparent to other links and to the protocol layer.

**Summary of Invention Paragraph:**

[0011] Some embodiments of the invention provide a computer system including a plurality of processor clusters interconnected by a plurality of point-to-point inter-cluster links, each processor cluster comprising nodes including a plurality of local processors and an interconnection controller interconnected by a plurality of point-to-point intra-cluster links, communications within a cluster being made via an intra-cluster protocol that uses intra-cluster packets, wherein the interconnection controller in each cluster is operable to map locally-generated communications directed to others of the clusters to the point-to-point inter-cluster links and to map remotely-generated communications directed to the local nodes to the point-to-point intra-cluster links, communications between clusters being made via an inter-cluster protocol that uses inter-cluster packets, an inter-cluster packet encapsulating at least one intra-cluster packet, each interconnection controller configured to compute a first cyclic redundancy code check for, and to encode first cyclic redundancy code check data in, each inter-cluster packet transmitted on the point-to-point inter-cluster links.

**Summary of Invention Paragraph:**

[0013] The nodes may be configured to compute a cyclic redundancy code check on a transmission window that includes transmissions of multiple intra-cluster packets

on the point-to-point intra-cluster links. Each interconnection controller may be further configured to compute a cyclic redundancy code check on a transmission window that includes transmissions of multiple intra-cluster packets on the point-to-point intra-cluster links. Each interconnection controller may encode the cyclic redundancy code check data in a field reserved for a link layer of each inter-cluster packet transmitted on the point-to-point inter-cluster links.

Summary of Invention Paragraph:

[0023] Some aspects of the invention provide a computer-implemented method for detecting errors in a computer system comprising a plurality of clusters, each cluster including a plurality of local nodes and an interconnection controller interconnected by point-to-point intra-cluster links, communications between the local nodes and the interconnection controller made via an intra-cluster protocol using intra-cluster packets, the interconnection controller of each cluster interconnected by point-to-point inter-cluster links with the interconnection controller of other clusters. The computer-implemented method includes the following steps: forming an inter-cluster packet by encapsulating an intra-cluster packet; encoding a sequence identifier in the inter-cluster packet; calculating first cyclic redundancy code check data based only upon the inter-cluster packet; encoding the first cyclic redundancy code check data in the inter-cluster packet; and transmitting the inter-cluster packet from a first interconnection controller to a second interconnection controller on a point-to-point inter-cluster link. The encoding steps may include encoding in an area of the inter-cluster packet reserved for link layer information.

Summary of Invention Paragraph:

[0025] Some embodiments of the invention provide an apparatus for detecting errors in a computer system comprising a plurality of clusters, each cluster including a plurality of local nodes and an interconnection controller interconnected by point-to-point intra-cluster links, communications between the local nodes and the interconnection controller made via an intra-cluster protocol using intra-cluster packets, the interconnection controller of each cluster interconnected by point-to-point inter-cluster links with the interconnection controller of other clusters. The apparatus includes: a device for forming an inter-cluster packet by encapsulating an intra-cluster packet; a device for encoding a sequence identifier in the inter-cluster packet; a device for calculating first cyclic redundancy code check data based only upon the inter-cluster packet; a device for encoding the first cyclic redundancy code check data in the inter-cluster packet; and a device for transmitting the inter-cluster packet from a first interconnection controller to a second interconnection controller on a point-to-point inter-cluster link.

CLAIMS:

1. A computer system comprising a plurality of processor clusters interconnected by a plurality of point-to-point inter-cluster links, each processor cluster comprising nodes including a plurality of local processors and an interconnection controller interconnected by a plurality of point-to-point intra-cluster links, communications within a cluster being made via an intra-cluster protocol that uses intra-cluster packets, wherein the interconnection controller in each cluster is operable to map locally-generated communications directed to others of the clusters to the point-to-point inter-cluster links and to map remotely-generated communications directed to the local nodes to the point-to-point intra-cluster links, communications between clusters being made via an inter-cluster protocol that uses inter-cluster packets, an inter-cluster packet encapsulating at least one intra-cluster packet, each interconnection controller configured to compute a first cyclic redundancy code check for, and to encode first cyclic redundancy code check data in, each inter-cluster packet transmitted on the point-to-point inter-cluster links.
26. A computer-implemented method for detecting errors in a computer system

comprising a plurality of clusters, each cluster including a plurality of local nodes and an interconnection controller interconnected by point-to-point intra-cluster links, communications between the local nodes and the interconnection controller made via an intra-cluster protocol using intra-cluster packets, the interconnection controller of each cluster interconnected by point-to-point inter-cluster links with the interconnection controller of other clusters, the computer-implemented method comprising: forming an inter-cluster packet by encapsulating an intra-cluster packet; encoding a sequence identifier in the inter-cluster packet; calculating first cyclic redundancy code check data based only upon the inter-cluster packet; encoding the first cyclic redundancy code check data in the inter-cluster packet; and transmitting the inter-cluster packet from a first interconnection controller to a second interconnection controller on a point-to-point inter-cluster link.

30. An apparatus for detecting errors in a computer system comprising a plurality of clusters, each cluster including a plurality of local nodes and an interconnection controller interconnected by point-to-point intra-cluster links, communications between the local nodes and the interconnection controller made via an intra-cluster protocol using intra-cluster packets, the interconnection controller of each cluster interconnected by point-to-point inter-cluster links with the interconnection controller of other clusters, the apparatus comprising: means for forming an inter-cluster packet by encapsulating an intra-cluster packet; means for encoding a sequence identifier in the inter-cluster packet; means for calculating first cyclic redundancy code check data based only upon the inter-cluster packet; means for encoding the first cyclic redundancy code check data in the inter-cluster packet; and means for transmitting the inter-cluster packet from a first interconnection controller to a second interconnection controller on a point-to-point inter-cluster link.

[Previous Doc](#)[Next Doc](#)[Go to Doc#](#)

[First Hit](#)[Previous Doc](#)[Next Doc](#)[Go to Doc#](#) [Generate Collection](#) [Print](#)

L7: Entry 1 of 13

File: PGPB

Nov 3, 2005

DOCUMENT-IDENTIFIER: US 20050243531 A1

TITLE: Interposer device

Summary of Invention Paragraph:

[0002] In some multi-processor systems, local nodes (including processors, input/output (I/O) devices, etc.) are directly connected to each other through a plurality of point-to-point intra-cluster links to form a cluster of two or more processors. The point-to-point links significantly increase the bandwidth for coprocessing and multiprocessing functions.

[Previous Doc](#)[Next Doc](#)[Go to Doc#](#)



US 20050034039A1

(19) United States

(12) Patent Application Publication (10) Pub. No.: US 2005/0034039 A1  
Prasad et al. (43) Pub. Date: Feb. 10, 2005

(54) METHODS AND DEVICES FOR INJECTING  
COMMANDS IN SYSTEMS HAVING  
MULTIPLE MULTI-PROCESSOR CLUSTERS

(75) Inventors: Guru Prasad, Austin, TX (US);  
David Brian Glasco, Austin, TX (US);  
Rajesh Kota, Austin, TX (US); Scott  
Diesing, Round Rock, TX (US)

Correspondence Address:  
BEYER WEAVER & THOMAS LLP  
P.O. BOX 778  
BERKELEY, CA 94704-0778 (US)

(73) Assignee: Newsys, Inc.

(21) Appl. No.: 10/635,700

(22) Filed: Aug. 5, 2003

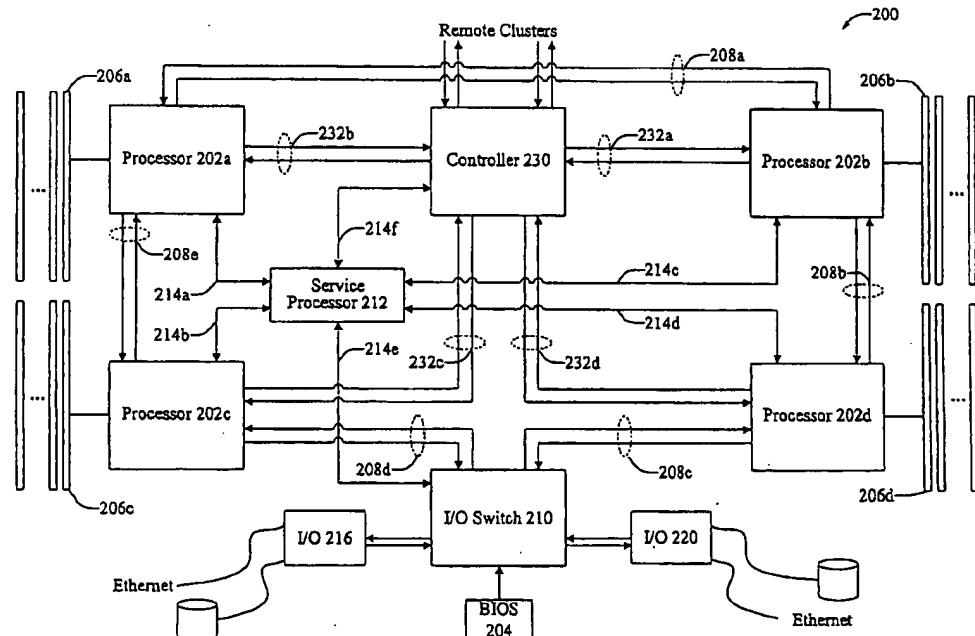
Publication Classification

(51) Int. Cl.<sup>7</sup> ..... G01R 31/28

(52) U.S. Cl. ..... 714/727

(57) ABSTRACT

Techniques and devices are provided for injecting transactions within computer systems having a plurality of multi-processor clusters. Each cluster includes a plurality of nodes, including processors, a service processor and an interconnection controller interconnected by point-to-point intra-cluster links. The processors and the interconnection controller in each cluster make transactions via an intra-cluster transaction protocol. Inter-cluster links are formed between interconnection controllers of different clusters. Each of the processors and the interconnection controller in a cluster has a test interface for communicating with the service processor. The service processor is configured to make an injected transaction according to the intra-cluster transaction protocol via one of the test interfaces. In preferred embodiments, the service processor is configured to make an injected transaction according to the intra-cluster transaction protocol via a test interface of an interconnection controller in the same cluster.



[First Hit](#) [Fwd Refs](#)[Previous Doc](#) [Next Doc](#) [Go to Doc#](#) [Generate Collection](#) [Print](#)

L11: Entry 2 of 3

File: USPT

Apr 6, 2004

DOCUMENT-IDENTIFIER: US 6718552 B1

**\*\* See image for Certificate of Correction \*\***

TITLE: Network bandwidth optimization by dynamic channel allocation

Detailed Description Text (21):

extended with a start of packet character (SOP), a routing header, and a cyclic redundancy check field. The SOP character is chosen from the ASI comma codes to ensure accurate detection of packets in the link. The routing header provides for a number of different routing possibilities including point to point and point to multi-point routing. A sixteen bit CRC is used for error detection on the rings 206, 208.

[Previous Doc](#) [Next Doc](#) [Go to Doc#](#)

[First Hit](#) [Fwd Refs](#)[Previous Doc](#) [Next Doc](#) [Go to Doc#](#)

End of Result Set

 [Generate Collection](#)  [Print](#)

L11: Entry 3 of 3

File: USPT

Aug 1, 2000

DOCUMENT-IDENTIFIER: US 6097707 A

TITLE: Adaptive digital wireless communications network apparatus and process

Brief Summary Text (31):

Multi-point digital wireless communications networks are also well known. In one system to Gilbert U.S. Pat. No. 5,297,144 a non-contention based, single optical-infrared channel, star configuration network using a central station to control access of multiple remote stations in a cyclic, synchronized, TDMA process is described. This patent describes a reservation period and a polling period protocol during which reserving stations from the first period are polled sequentially for data transfer in the second period. The system does not describe a wireless radio frequency link and suffers from less than optimal utilization of the channel.

[Previous Doc](#) [Next Doc](#) [Go to Doc#](#)

[First Hit](#) [Fwd Refs](#)[Previous Doc](#) [Next Doc](#) [Go to Doc#](#) [Generate Collection](#)  [Print](#)

L18: Entry 14 of 18

File: USPT

Feb 19, 2002

DOCUMENT-IDENTIFIER: US 6349091 B1

TITLE: Method and apparatus for controlling communication links between network nodes to reduce communication protocol overhead traffic

Brief Summary Text (6):

The intranet protocol is preferably a link-state type of routing protocol that is implemented on the backbone network. The cluster head nodes each include a database that is synchronized with other cluster head node databases by transference or flooding of Link-State Advertisement (LSA) packets in accordance with the protocol. These databases include information enabling the cluster head nodes to determine appropriate paths for routing messages through the network, while the LSA packets provide information to update the databases. Since the two-tier network is not a broadcast type network, LSA and other routing control packets are sent by a cluster head node as point-to-point messages to each of its cluster head neighbors.

Detailed Description Text (4):

The arrangement of nodes 10 within clusters 12 and the designation of cluster head nodes 14 are dynamically determined by an intranet protocol that handles routing of packets within network 2. The intranet protocol is preferably a link-state type routing protocol that is implemented on backbone network 16. Cluster head nodes 14 each maintain a database including information enabling that cluster head to determine appropriate paths for routing messages through the network. The information typically relates to links between the various network nodes. The cluster head databases are synchronized in accordance with the protocol by transference of link-State Advertisement (LSA) packets between cluster head nodes that provide network information to update the databases. These packets are conventionally transmitted to each neighboring network node via plural point-to-point messages (e.g., messages from a source node to a specific destination network node) in response to changes in network topology, an external network connected to network 2 or other modifications to the network facilitating changes in a node database. When a database update packet is received, a point-to-point acknowledgment (ACK) packet is commonly transmitted to the source node from the destination node to indicate packet reception. In addition, each node (e.g., cluster head and member nodes) periodically broadcasts a beacon type or node status packet in accordance with the intranet protocol. This packet basically advertises the presence of a node within the network. The rate of transmission of a node status packet by a cluster head node is significantly greater than the transmission rate for that packet by a cluster member node since cluster head node status packets are typically utilized for "keep alive" and neighbor discovery purposes.

[Previous Doc](#) [Next Doc](#) [Go to Doc#](#)



US 20050034049A1

(19) United States

(12) Patent Application Publication (10) Pub. No.: US 2005/0034049 A1  
(43) Pub. Date: Feb. 10, 2005

(54) COMMUNICATION BETWEEN  
MULTI-PROCESSOR CLUSTERS OF  
MULTI-CLUSTER COMPUTER SYSTEMS

(75) Inventors: Shashank Nemawarkar, Austin, TX  
(US); Rajesh Kota, Austin, TX (US);  
Guru Prasad, Austin, TX (US); Carl  
Zeitler, Tomball, TX (US); David B.  
Glasco, Austin, TX (US)

Correspondence Address:  
BEYER WEAVER & THOMAS LLP  
P.O. BOX 778  
BERKELEY, CA 94704-0778 (US)

(73) Assignee: Newisys, Inc.

(21) Appl. No.: 10/635,884

(22) Filed: Aug. 5, 2003

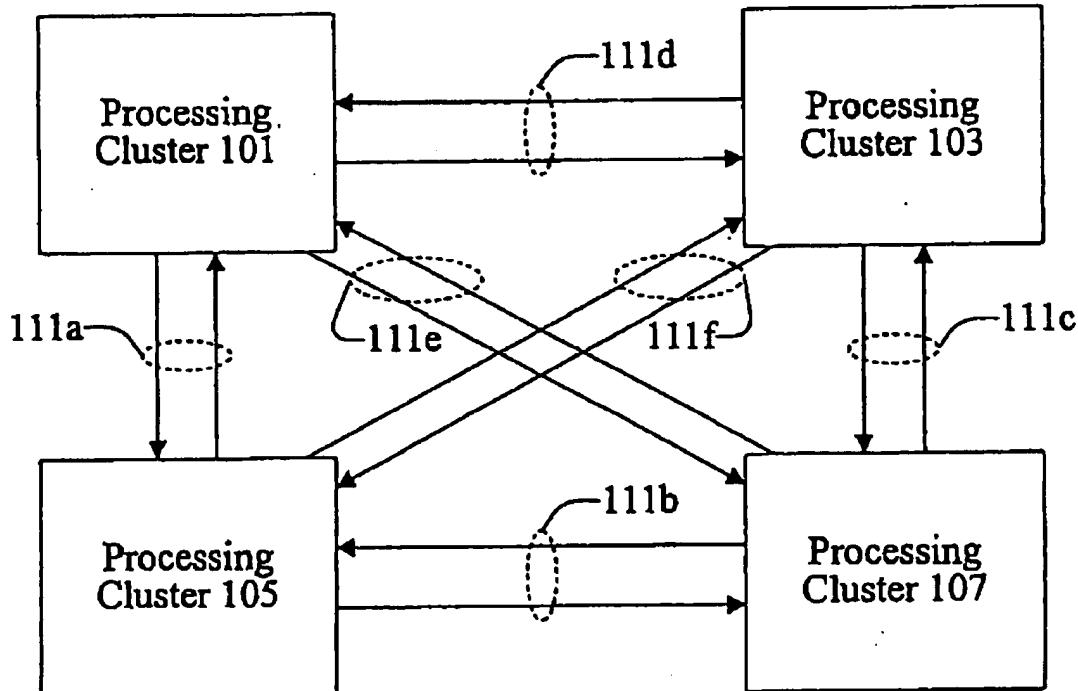
Publication Classification

(51) Int. Cl.<sup>7</sup> ..... H03M 13/00

(52) U.S. Cl. ..... 714/758

(57) ABSTRACT

Improved techniques are provided for detecting and correcting errors and skew in inter-cluster communications within computer systems having a plurality of multi-processor clusters. The local nodes of each cluster include a plurality of processors and an interconnection controller. Intra-cluster links are formed between the local nodes, including the interconnection controller, within a cluster. Inter-cluster links are formed between interconnection controllers of different clusters. Intra-cluster packets may be serialized and encapsulated as inter-cluster packets for transmission on inter-cluster links, preferably with link-layer encapsulation. Each inter-cluster packet may include a sequence identifier and error information computed for that packet. Clock data may be embedded in symbols sent on each bit lane of the inter-cluster links. Copies of transmitted inter-cluster packets may be stored until an acknowledgement is received. The use of inter-cluster packets on an inter-cluster link is preferably transparent to other links and to the protocol layer.





US 20050034048A1

(19) United States

(12) Patent Application Publication (10) Pub. No.: US 2005/0034048 A1  
Nemawarkar et al. (43) Pub. Date: Feb. 10, 2005

(54) RELIABLE COMMUNICATION BETWEEN  
MULTI-PROCESSOR CLUSTERS OF  
MULTI-CLUSTER COMPUTER SYSTEMS

(75) Inventors: Shashank Nemawarkar, Austin, TX  
(US); Rajesh Kota, Austin, TX (US);  
Guru Prasad, Austin, TX (US); Carl  
Zeitler, Tomball, TX (US); David B.  
Glasco, Austin, TX (US)

Correspondence Address:  
BEYER WEAVER & THOMAS LLP  
P.O. BOX 778  
BERKELEY, CA 94704-0778 (US)

(73) Assignee: Newisys, Inc.

(21) Appl. No.: 10/635,793

(22) Filed: Aug. 5, 2003

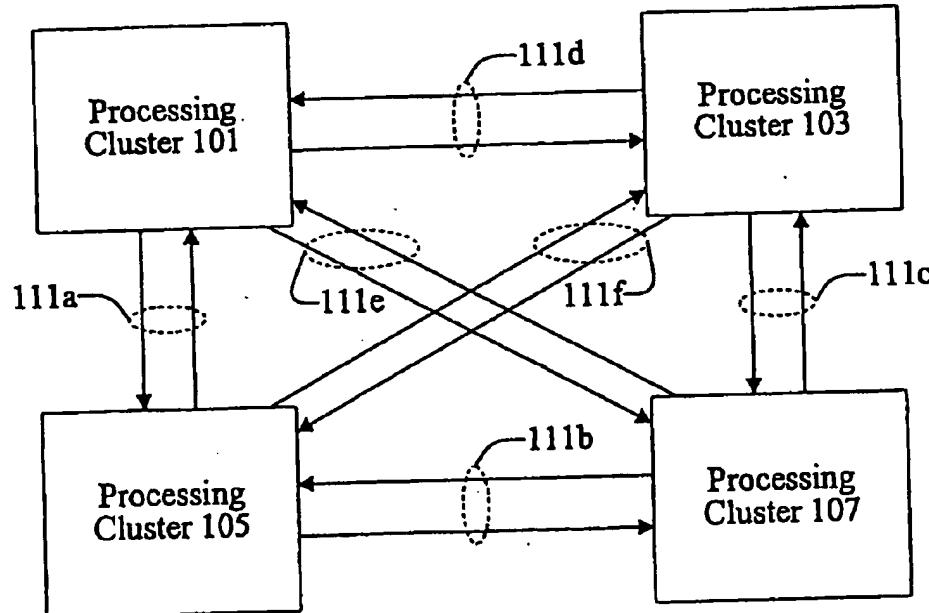
Publication Classification

(51) Int. Cl.<sup>7</sup> ..... H03M 13/00

(52) U.S. Cl. ..... 714/758

(57) ABSTRACT

Improved techniques are provided for detecting and correcting errors and skew in inter-cluster communications within computer systems having a plurality of multi-processor clusters. The local nodes of each cluster include a plurality of processors and an interconnection controller. Intra-cluster links are formed between the local nodes, including the interconnection controller, within a cluster. Inter-cluster links are formed between interconnection controllers of different clusters. Intra-cluster packets may be serialized and encapsulated as inter-cluster packets for transmission on inter-cluster links, preferably with link-layer encapsulation. Each inter-cluster packet may include a sequence identifier and error information computed for that packet. Clock data may be embedded in symbols sent on each bit lane of the inter-cluster links. Copies of transmitted inter-cluster packets may be stored until an acknowledgement is received. The use of inter-cluster packets on an inter-cluster link is preferably transparent to other links and to the protocol layer.





US 20050034007A1

(19) **United States**

**(12) Patent Application Publication  
Nemawarkar et al.**

(10) Pub. No.: US 2005/0034007 A1

(43) Pub. Date: Feb. 10, 2005

(54) SYNCHRONIZED COMMUNICATION  
BETWEEN MULTI-PROCESSOR CLUSTERS  
OF MULTI-CLUSTER COMPUTER SYSTEMS

(75) Inventors: **Shashank Nemawarkar, Austin, TX (US); Rajesh Kota, Austin, TX (US); Guru Prasad, Austin, TX (US); Carl Zeitler, Tomball, TX (US); David B. Glasco, Austin, TX (US)**

**Correspondence Address:**  
**BEYER WEAVER & THOMAS LLP**  
**P.O. BOX 778**  
**BERKELEY, CA 94704-0778 (US)**

(73) Assignee: Newisys, Inc.

(21) Appl. No.: 10/635,705

(22) Filed: Aug. 5, 2003

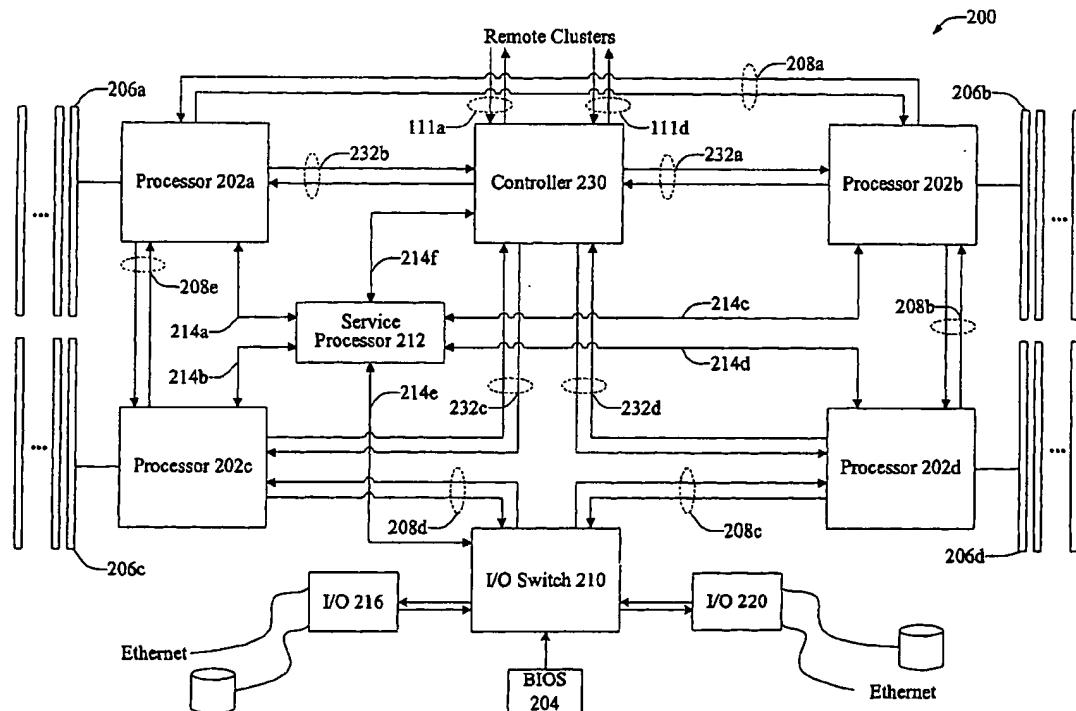
#### Publication Classification

(51) Int. Cl.<sup>7</sup> ..... G06F 1/04

(52) U.S. Cl. .... 713/500

## ABSTRACT

Improved techniques are provided for detecting and correcting errors and skew in inter-cluster communications within computer systems having a plurality of multi-processor clusters. The local nodes of each cluster include a plurality of processors and an interconnection controller. Intra-cluster links are formed between the local nodes, including the interconnection controller, within a cluster. Inter-cluster links are formed between interconnection controllers of different clusters. Intra-cluster packets may be serialized and encapsulated as inter-cluster packets for transmission on inter-cluster links, preferably with link-layer encapsulation. Each inter-cluster packet may include a sequence identifier and error information computed for that packet. Clock data may be embedded in symbols sent on each bit lane of the inter-cluster links. Copies of transmitted inter-cluster packets may be stored until an acknowledgement is received. The use of inter-cluster packets on an inter-cluster link is preferably transparent to other links and to the protocol layer.





US 20050034033A1

(19) United States

(12) Patent Application Publication (10) Pub. No.: US 2005/0034033 A1  
Nemawarkar et al. (43) Pub. Date: Feb. 10, 2005

(54) COMMUNICATION BETWEEN AND WITHIN  
MULTI-PROCESSOR CLUSTERS OF  
MULTI-CLUSTER COMPUTER SYSTEMS

(75) Inventors: Shashank Nemawarkar, Austin, TX  
(US); Rajesh Kota, Austin, TX (US);  
Guru Prasad, Austin, TX (US); Carl  
Zeitler, Tomball, TX (US); David B.  
Glasco, Austin, TX (US)

Correspondence Address:  
BEYER WEAVER & THOMAS LLP  
P.O. BOX 778  
BERKELEY, CA 94704-0778 (US)

(73) Assignee: Newisys, Inc.

(21) Appl. No.: 10/635,744

(22) Filed: Aug. 5, 2003

Publication Classification

(51) Int. Cl.<sup>7</sup> ..... G06F 11/00

(52) U.S. Cl. ..... 714/52

(57) ABSTRACT

Improved techniques are provided for detecting and correcting errors and skew in inter-cluster communications within computer systems having a plurality of multi-processor clusters. The local nodes of each cluster include a plurality of processors and an interconnection controller. Intra-cluster links are formed between the local nodes, including the interconnection controller, within a cluster. Inter-cluster links are formed between interconnection controllers of different clusters. Intra-cluster packets may be serialized and encapsulated as inter-cluster packets for transmission on inter-cluster links, preferably with link-layer encapsulation. Each inter-cluster packet may include a sequence identifier and error information computed for that packet. Clock data may be embedded in symbols sent on each bit lane of the inter-cluster links. Copies of transmitted inter-cluster packets may be stored until an acknowledgement is received. The use of inter-cluster packets on an inter-cluster link is preferably transparent to other links and to the protocol layer.

